Optimizing X-ray optical prescriptions for wide-field applications

R. F. Elsner, S. L. O'Dell, B. D. Ramsey, and M. C. Weisskopf, NASA/MSFC

X-ray telescopes with spatial resolution optimized over the field-of-view (FOV) are of special interest for missions, such as WFXT, focused on moderately deep and deep surveys of the X-ray sky, and for solar X-ray observations. Here we report on the present status of an on-going study of the properties of Wolter I and polynomial grazing incidence designs, with a view to gaining deeper insight into their properties and simplifying the design process. With these goals in mind, we present some results in the complementary topics of (1) properties of Wolter I X-ray optics and (2) polynomial X-ray optic ray tracing. Of crucial importance for the design of wide-field X-ray optics is the optimization criteria. Here we have adopted the minimization of a merit function, M, which measures the spatial resolution averaged over the FOV:

Merit Function:

$$M = \frac{\int_0^{2\pi} d\phi \int_0^{\theta_{FOV}} d\theta\, \theta\, w(\theta)\sigma^2(\theta,\phi)}{\int_0^{\pi/4} d\phi \int_0^{\theta_{FOV}} d\theta\, \theta\, w(\theta)}$$

where $w(\theta,)$ is a weighting function and

$$\sigma^2(\theta,\phi) = \sum_{(x,y,z)} \left[<(x,y,z)^2> - <(x,y,z)>^2 \right]$$

is the spatial variance for a point source on the sky at polar and azimuthal off-axis angles (θ,φ).

Single Shell:

For a single Wolter I mirror shell s, the spatial variance may be written as

$$\sigma_s^2(\theta,\phi) = a_s + 2\,b_s\,\delta z_s + c_s\,\delta z_s^2 + 2\,d_s \tan\theta_{tilt} + 2\,e_s\,\delta z_s \tan\theta_{tilt} + f_s \tan^2\theta_{tilt}$$
$$a_s = \sum_{(x,y)} \left[<(x,y)^2>_{0,s} - <(x,y)>_{0,s}^2 \right]$$
$$b_s = \sum_{(x,y)} \left[<(x,y)(k_{(x,y)}/k_z)>_{0,s} - <(x,y)>_{0,s} <(k_{(x,y)}/k_z)>_{0,s} \right]$$
$$c_s = \sum_{(x,y)} \left[<(k_{(x,y)}/k_z)^2>_{0,s} - <k_{(x,y)}/k_z>^2 \right]$$

The angle brackets around a quantity q, denote an average over a set of exit rays from a Monte-Carlo ray trace, and the subscript *0* denotes evaluation in the flat plane perpendicular to the optical axis at the nominal on-axis focus. The variable δz_s is the displacement of the mirror shell along the optical axis from its nominal position. The variable θ_{tilt} is the angle with by which a detector plane with corner at the optical axis is tilted (see Figure 1). Expressions for d_s, e_s and f_s are available on request. Using the results from extensive Monte-Carlo ray traces (see Figure 2), we have devised analytic trail functions for the coefficients a_s, b_s, c_s, d_s, e_s and f_s:

$$a_{s,trial}(\theta) = \left[a_0 = \left(\frac{2\,\mu_{a,0}\,\ell}{\tan 4\alpha_0}\right)^2 \tan^4\theta\;\; g(\theta,\varsigma_a,\xi_a) \right] + \left[a_{coma} = \left(\frac{\tan 4\alpha_0}{2}\right)^4 \tan^2\theta \right]$$
$$b_{s,trial}(\theta) = 2\,\mu_b\,\ell \tan^2\theta\;\; g(\theta,\varsigma_b,\xi_b)$$
$$c_{s,trial}(\theta) = \left[\mu_c \tan 4\alpha_{0,s}\;\; g(\theta,\varsigma_c,\xi_c)\right]^2$$
$$g(\theta,\varsigma,\xi) = 1 + \varsigma\tan\theta + \xi\tan^2\theta$$

Expressions for $d_{s,trial}$, $e_{s,trial}$ and $f_{s,trial}$ are available on request. Because of the azimuthal asymmetry introduced by tilting the detectors, these coefficients depend on azimuthal off-axis angle as well as polar off-axis angle. Integrating to evaluate the merit function M, and then minimizing M, we find:

$$M = a_{s,M} + 2\,b_{s,M}\,\delta z_s + c_{s,M}\,\delta z_s^2 + 2\,d_{s,M}\tan\theta_{tilt} + 2\,e_{s,M}\,\delta z_s\tan\theta_{tilt} + f_{s,M}\tan^2\theta_{tilt}$$
$$\tan\theta_{tilt} = \left(\frac{b_{s,M}\,e_{s,M} - c_{s,M}\,d_{s,M}}{c_{s,M}\,f_{s,M} - e_{s,M}^2}\right)$$
$$\delta z_s = \left(\frac{d_{s,M}\,e_{s,M} - b_{s,M}\,d_{s,M}}{c_{s,M}\,f_{s,M} - e_{s,M}^2}\right)$$

Nested Shells:

For a set of S nested mirror shells, we find the spatial variance is the sum of two terms. The first is a sum over the spatial variances of the individual shells, weighted by their effective areas. The second is a weighted sum over a kind of variance of the spatial means for the individual shells. The second term can be viewed as arising from the fact that the best focal surfaces of the different shells do not coincide with each other or with the best focal surface for the nested set. This result means that the shell parameters must be optimized simultaneously rather than individually. These results for nested shells and the expressions below are valid for any mirror surface prescription.

$$\sigma^2(\theta,\phi) = \sigma_1^2(\theta,\phi) + \sigma_2^2(\theta,\phi)$$
$$\sigma_1^2(\theta,\phi) = \sum_{s=1}^{S}\left(\frac{n_s - 1}{N-1}\right)\sigma_s^2(\theta,\phi)$$
$$\sigma_2^2(\theta,\phi) = \left(\frac{N}{N-1}\right)\sum_{(x,y,z)}\left[\sum_{s=1}^{S}\left(\frac{n_s}{N}\right)<(x,y,z)>^2 - \left(\sum_{s=1}^{S}\left(\frac{n_s}{N}\right)<(x,y,z)>\right)^2\right]$$
$$N = \sum_{s=1}^{S} n_s$$

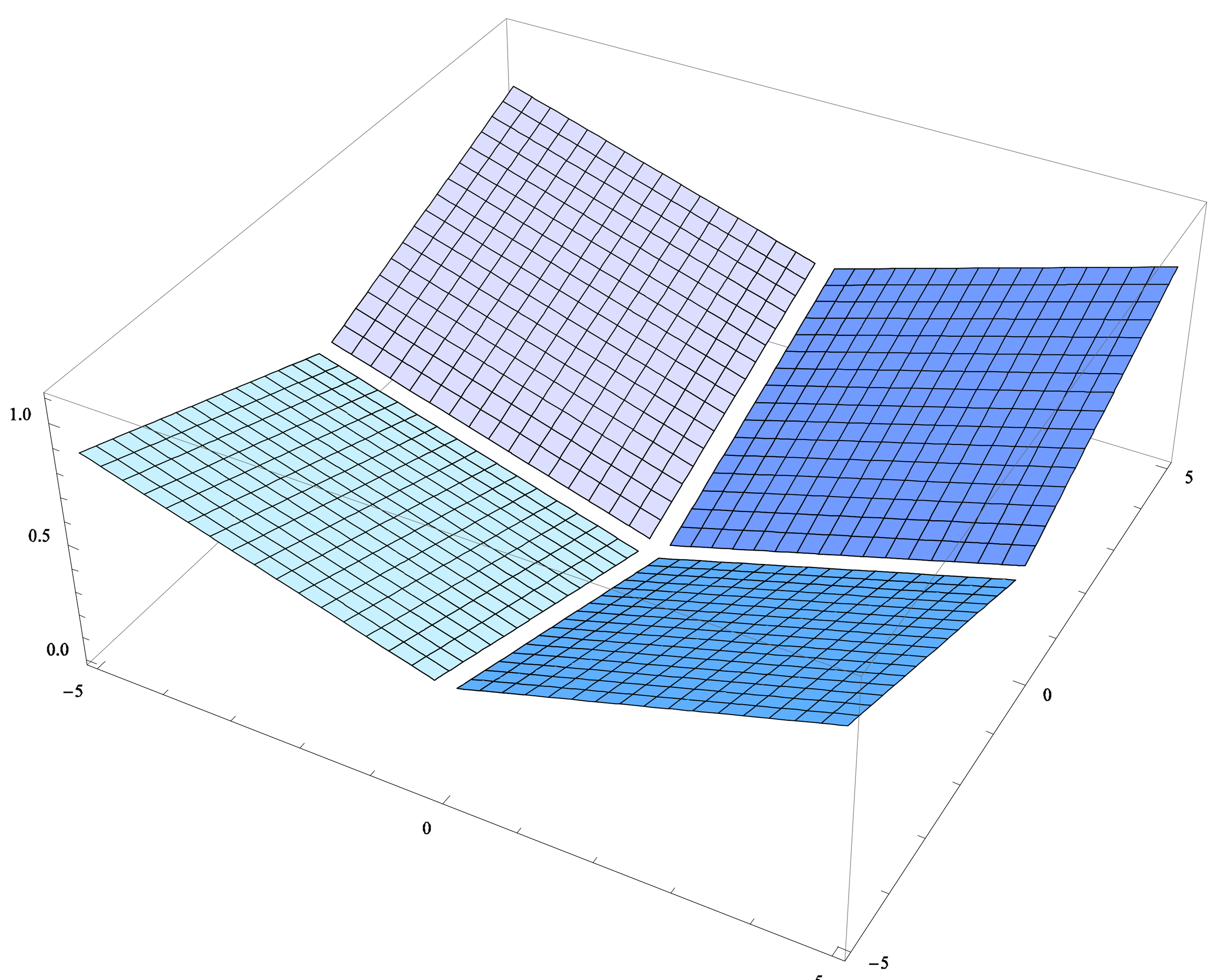

Figure 1. Detector geometry showing 4 flat tilted detectors, one per quadrant. The detectors are tilted up from a flat plane perpendicular to the optical axis by an amount Θ_{tilt}.

Ray Tracing Polynomial X-ray Optics:

In 1992, Burrows, Burg, and Giacconi showed how adding higher order polynomial terms to Wolter I prescriptions, and hence giving up some on-axis spatial resolution, can lead to mirror surface prescriptions with improved spatial resolution over a wide FOV. Optimizing a nested array of polynomial optics involves many Monte-Carlo ray traces to cover parameter space, and complex methods of finding the optimum design. We have devised a method, valid when the polynomial coefficients are sufficiently small, for ray tracing polynomial optics keeping the polynomial coefficients in symbolic form. The method treats a polynomial optic as a perturbation on an underlying Wolter I optic:

$$r_s^2(z) = r_{0,s}^2\left[1 + 2\,A_s\left(\frac{z}{r_{0,s}}\right) + B_s\left(\frac{z}{r_{0,s}}\right)^2 + u_{a,s}\left(\frac{z}{r_{0,s}}\right)^2 + u_{b,s}\left(\frac{z}{r_{0,s}}\right)^3\right]$$

where $u_{a,s}$ and $u_{b,s}$ are the polynomial coefficients and differ for the primary, *P*, and secondary, *S*, mirror elements (higher order polynomial terms may also be included), as do A_s and B_s. For an underlying Wolter I optic we have:

$$P:\quad A_s = \tan\alpha_{0,s} \qquad B_s = 0$$
$$S:\quad A_s = \tan 3\alpha_{0,s} \qquad B_s = h(\alpha_{0,s})\tan^2 3\alpha_{0,s}$$
$$\alpha_{0,s} = \left(\frac{1}{4}\right)\tan^{-1}\left(\frac{r_{0,s}}{f}\right)$$
$$h(\alpha_{0,s}) = 1 - \frac{1}{[1 + 2\cos(2\alpha_{0,s})]^2}$$

The angle $\alpha_{0,s}$ is the graze angle for an on-axis ray at the intersection plane.

Our method involves writing ray position components (x,y,z) and direction vectors (k_x,k_y,k_z) in the form:

$$\lambda = \lambda_{00} + \vec{u}\cdot\vec{\lambda}_0 + \vec{u}\cdot\overline{\overline{\lambda}}\cdot\vec{u}$$

where $\vec{u} = \left(u_{a,s,P},\; u_{b,s,P},\; u_{a,s,S},\; u_{b,s,S}\right)$

In this notation, the mirror prescription becomes:

$$r_s^2(z) = r_{0,s}^2\left[1 + 2\,A_s\left(\frac{z}{r_{0,s}}\right) + B_s\left(\frac{z}{r_{0,s}}\right)^2 + \left(\frac{z}{r_{0,s}}\right)^2\vec{u}\cdot\vec{\varsigma}_{0,s}\right]$$

$$\vec{\varsigma}_{0,s,P} = \left(1,\frac{z}{r_{0,s}},0,0\right) \qquad \vec{\varsigma}_{0,s,S} = \left(0,0,1,\frac{z}{r_{0,s}}\right)$$

As long as the polynomial coefficients are sufficiently small, various operations such as addition, subtraction, multiplication, division, taking of square roots, etc. are evaluated by expansions to second order in the polynomial coefficients. Then all the tasks required for a ray trace can be accomplished to the appropriate order. These tasks are (1) populating the entrance aperture with rays (in position, the incident direction is assumed), (2) finding intersections with mirror segment surfaces, (3) calculating unit normals to those surfaces, including deviations due to non-ideal surfaces, (4) determining the direction vector of the reflected ray, and (5) taking account of any obstruction by the next innermost shell.

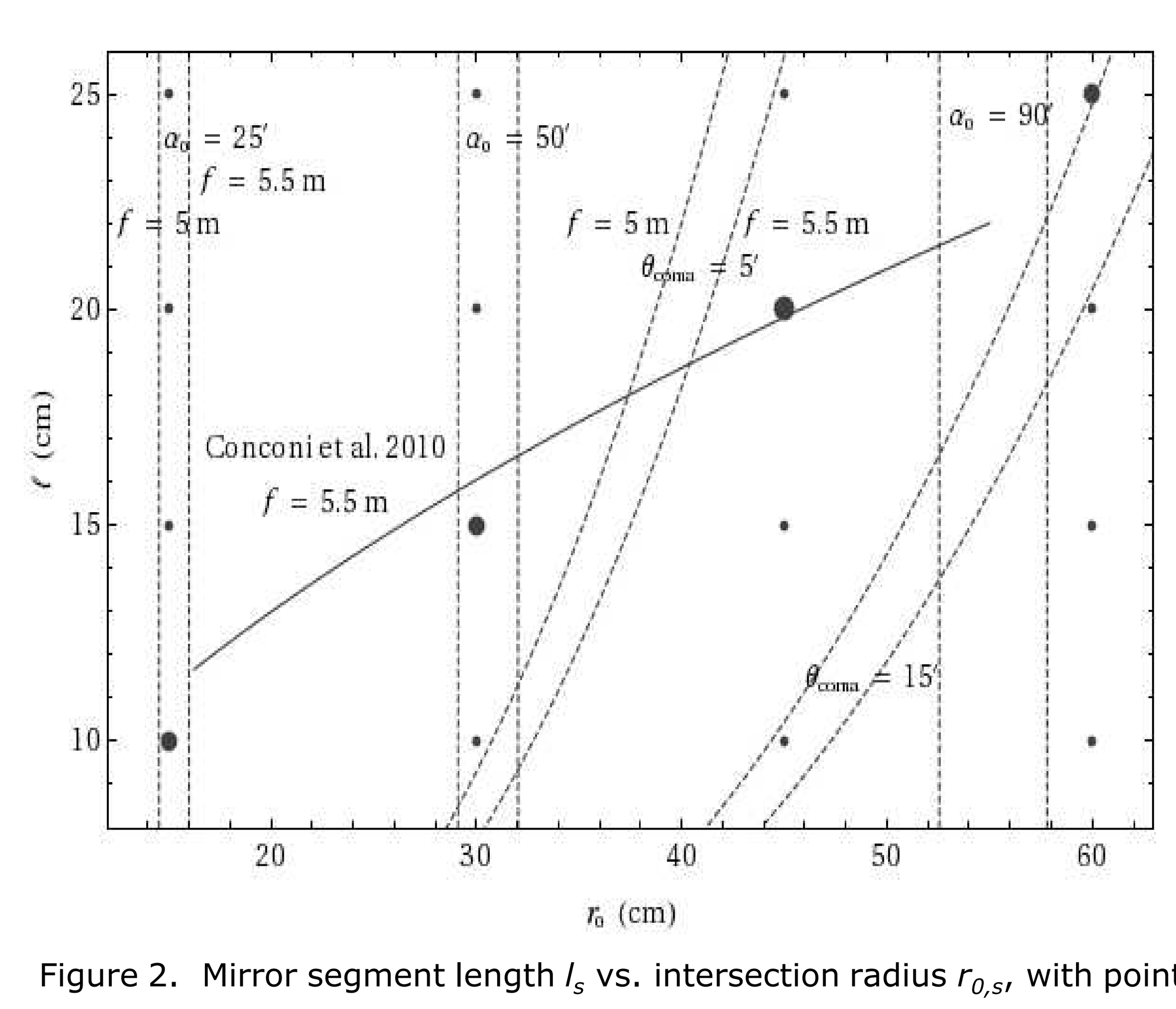

Figure 2. Mirror segment length l_s vs. intersection radius $r_{0,s}$, with points showing locations in the $(l_s, r_{0,s})$ plane of Monte-Carlo ray traces with 50,000 incident rays for f = 5.5 m. For the largest dot, we also carried out ray traces with 100,000 incident rays for f = 5.5 m. For the largest and mid-size dots, we also carried out ray traces for f = 4.5, 5.0 and 6.0 with 50,000 incident rays. The solid curve shows the l_s vs. $r_{0,s}$ relation for the wide-field design of Conconi et al. (2010). Curved dotted lines track the angle θ_{coma} for which $a_{coma} = a_0$.

Optimizing X-ray optical prescriptions for wide-field applications

R. F. Elsner, S. L. O'Dell, B. D. Ramsey, and M. C. Weisskopf, NASA/MSFC

X-ray telescopes with spatial resolution optimized over the field-of-view (FOV) are of special interest for missions, such as WFXT, focused on moderately deep and deep surveys of the X-ray sky, and for solar X-ray observations. Here we report on the present status of an on-going study of the properties of Wolter I and polynomial grazing incidence designs, with a view to gaining deeper insight into their properties and simplifying the design process. With these goals in mind, we present some results in the complementary topics of (1) properties of Wolter I X-ray optics and (2) polynomial X-ray optic ray tracing. Of crucial importance for the design of wide-field X-ray optics is the optimization criteria. Here we have adopted the minimization of a merit function, M, which measures the spatial resolution averaged over the FOV.

Merit Function:

Of crucial importance for the design of wide-field X-ray optics is the optimization criteria. Here we have adopted the minimization of a merit function, *M*, which measures the spatial resolution averaged over the FOV:

$$M = \frac{\int_0^{2\pi} d\phi \int_0^{\theta_{FOV}} d\theta\, \theta\, w(\theta)\, \sigma^2(\theta,\phi)}{\int_0^{\pi/4} d\phi \int_0^{\theta_{FOV}} d\theta\, \theta\, w(\theta)}$$

where w(θ,) is a weighting function and

$$\sigma^2(\theta,\phi) = \sum_{(x,y,z)} \left[< (x,y,z)^2 > - < (x,y,z) >^2 \right]$$

is the spatial variance for a point source on the sky at polar and azimuthal off-axis angles (θ,φ).

Single Shell:

For a single Wolter I mirror shell *s*, the spatial variance may be written as

$$\sigma_s^2(\theta,\phi) = a_s + 2\,b_s\,\delta z_s + c_s\,\delta z_s^2 + 2\,d_s \tan\theta_{tilt} + 2\,e_s\,\delta z_s \tan\theta_{tilt} + f_s \tan^2\theta_{tilt}$$

$$a_s = \sum_{(x,y)} \left[< (x,y)^2 >_{0,s} - < (x,y) >_{0,s}^2 \right]$$

$$b_s = \sum_{(x,y)} \left[< (x,y)(k_{(x,y)}/k_z) >_{0,s} - < (x,y) >_{0,s} < (k_{(x,y)}/k_z) >_{0,s} \right]$$

$$c_s = \sum_{(x,y)} \left[< (k_{(x,y)}/k_z)^2 >_{0,s} - < k_{(x,y)}/k_z >^2 \right]$$

The angle brackets around a quantity q, denote an average over a set of exit rays from a Monte-Carlo ray trace, and the subscript *0* denotes evaluation in the flat plane perpendicular to the optical axis at the nominal on-axis focus. The variable δz_s is the displacement of the mirror shell along the optical axis from its nominal position. The variable θ_{tilt} is the angle with by which a detector plane with corner at the optical axis is tilted (see Figure 1). Expressions for d_s, e_s and f_s are available on request.

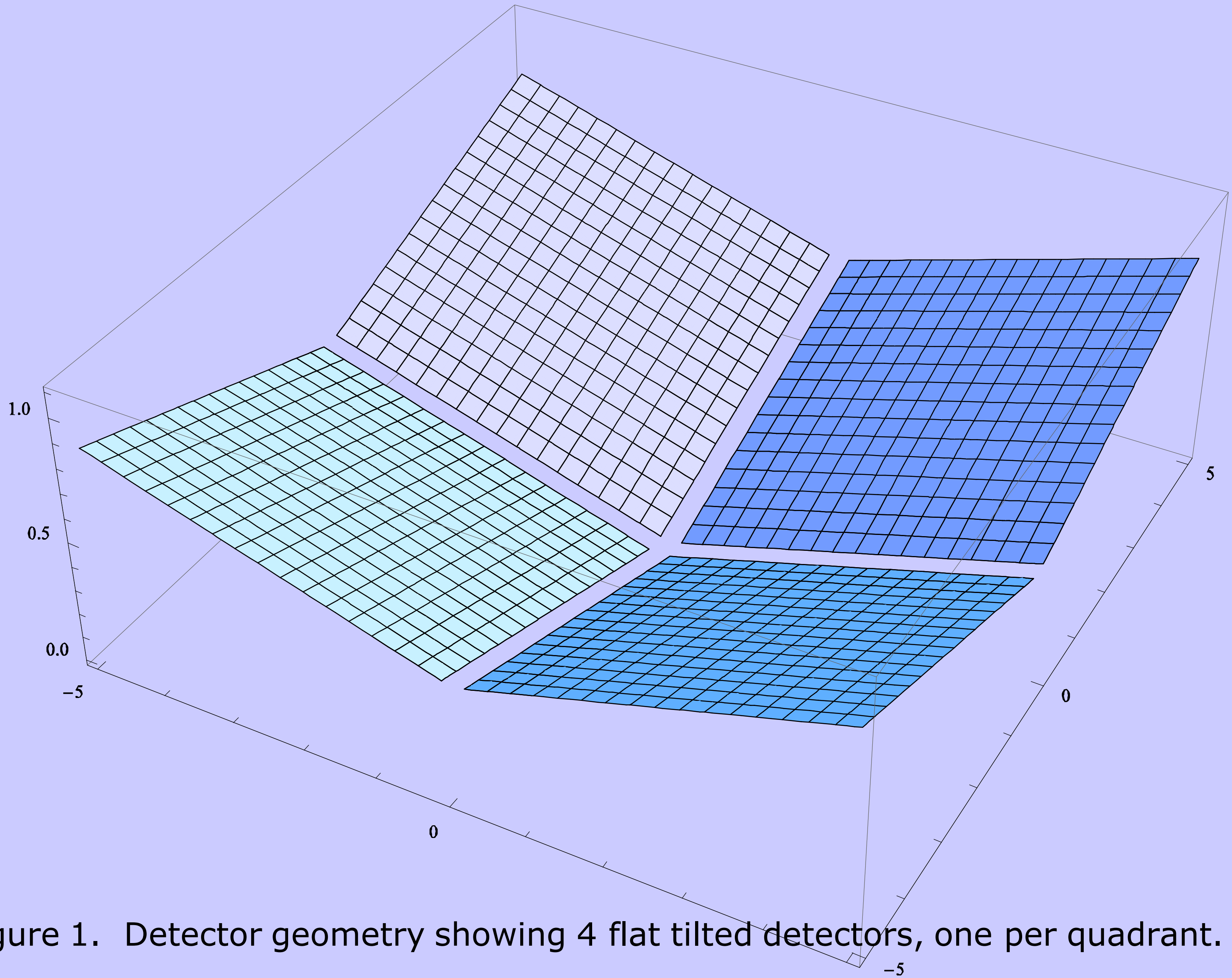

Figure 1. Detector geometry showing 4 flat tilted detectors, one per quadrant. The detectors are tilted up from a flat plane perpendicular to the optical axis by an amount Θ_{tilt}.

Using the results from extensive Monte-Carlo ray traces (see Figure 2), we have devised analytic trail functions for the coefficients a_s, b_s, c_s, d_s, e_s and f_s:

$$a_{s,trial}(\theta) = \left[a_0 = \left(\frac{2\,\mu_{a,0}\,\ell}{\tan 4\alpha_0} \right)^2 \tan^4\theta \;\; g(\theta,\varsigma_a,\xi_a) \right] + \left[a_{coma} = \left(\frac{\tan 4\alpha_0}{2} \right)^4 \tan^2\theta \right]$$

$$b_{s,trial}(\theta) = 2\,\mu_b\,\ell \tan^2\theta \;\; g(\theta,\varsigma_b,\xi_b)$$

$$c_{s,trial}(\theta) = \left[\mu_c \tan 4\alpha_{0,s} \;\; g(\theta,\varsigma_c,\xi_c) \right]^2$$

$$g(\theta,\varsigma,\xi) = 1 + \varsigma \tan\theta + \xi \tan^2\theta$$

Expressions for $d_{s,trial}$, $e_{s,trial}$ and $f_{s,trial}$ are available on request. Because of the azimuthal asymmetry introduced by tilting the detectors, these coefficients depend on azimuthal off-axis angle as well as polar off-axis angle.

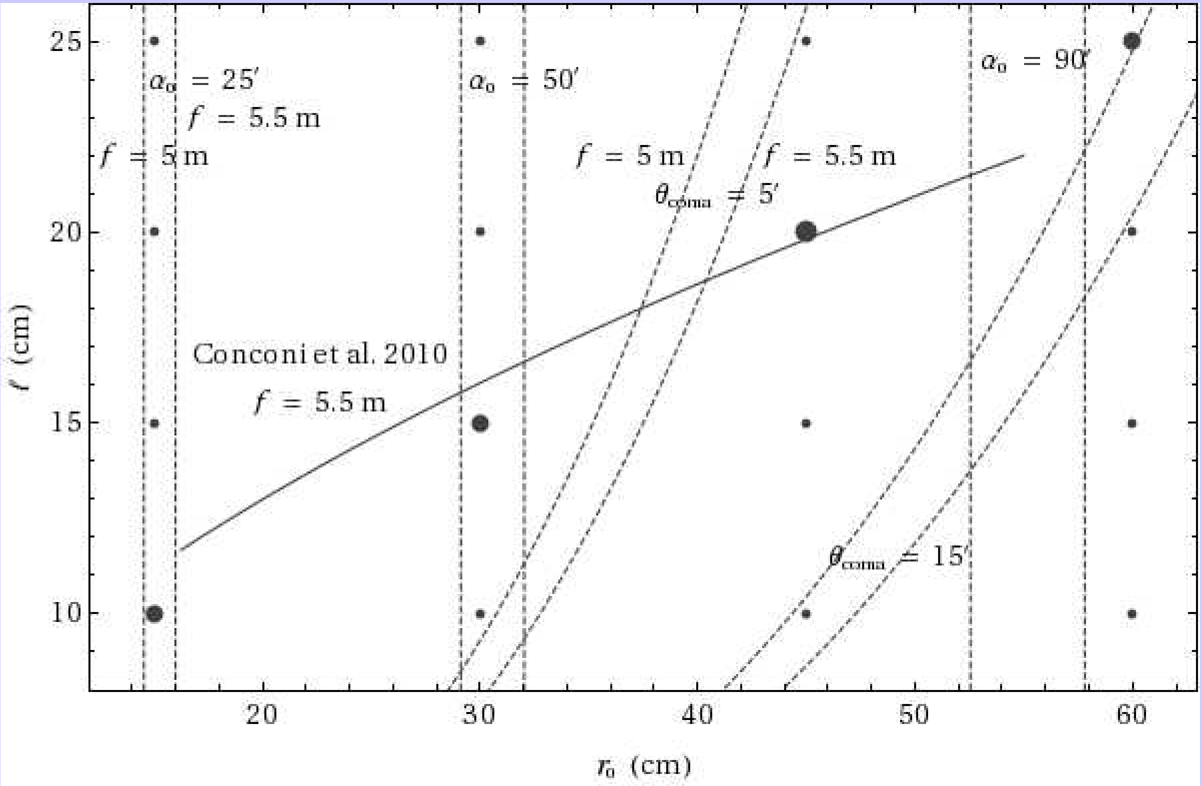

Figure 2. Mirror segment length l_s vs. intersection radius $r_{0,s}$, with points showing locations in the (l_s, $r_{0,s}$) plane of Monte-Carlo ray traces with 50,000 incident rays for f = 5.5 m. For the largest dot, we also carried out ray traces with 100,000 incident rays for f = 5.5 m. For the largest and mid-size dots, we also carried out ray traces for f = 4.5, 5.0 and 6.0 with 50,000 incident rays. The solid curve shows the l_s vs. $r_{0,s}$ relation for the wide-field design of Conconi et al. (2010). Curved dotted lines track the angle θ_{coma} for which a_{coma} = a_0.

Integrating to evaluate the merit function M, and then minimizing M, we find:

$$M = a_{s,M} + 2\, b_{s,M}\, \delta z_s + c_{s,M}\, \delta z_s^2 + 2\, d_{s,M} \tan\theta_{tilt} + 2\, e_{s,M}\, \delta z_s \tan\theta_{tilt} + f_{s,M} \tan^2\theta_{tilt}$$

$$\tan\theta_{tilt} = \left(\frac{b_{s,M}\, e_{s,M} - c_{s,M}\, d_{s,M}}{c_{s,M}\, f_{s,M} - e_{s,M}^2} \right)$$

$$\delta z_s = \left(\frac{d_{s,M}\, e_{s,M} - b_{s,M}\, d_{s,M}}{c_{s,M}\, f_{s,M} - e_{s,M}^2} \right)$$

Nested Shells:

For a set of S nested mirror shells, we find the spatial variance is the sum of two terms. The first is a sum over the spatial variances of the individual shells, weighted by their effective areas. The second is a weighted sum over a kind of variance of the spatial means for the individual shells. The second term can be viewed as arising from the fact that the best focal surfaces of the different shells do not coincide with each other or with the best focal surface for the nested set. This result means that the shell parameters must be optimized simultaneously rather than individually. These results for nested shells and the expressions below are valid for any mirror surface prescription.

$$\sigma^2(\theta,\phi) = \sigma_1^2(\theta,\phi) + \sigma_2^2(\theta,\phi)$$

$$\sigma_1^2(\theta,\phi) = \sum_{s=1}^{S}\left(\frac{n_s - 1}{N-1}\right)\sigma_s^2(\theta,\phi)$$

$$\sigma_2^2(\theta,\phi) = \left(\frac{N}{N-1}\right)\sum_{(x,y,z)}\left[\sum_{s=1}^{S}\left(\frac{n_s}{N}\right) < (x,y,z) >^2 - \left(\sum_{s=1}^{S}\left(\frac{n_s}{N}\right) < (x,y,z) >\right)^2\right]$$

$$N = \sum_{s=1}^{S} n_s$$

Ray Tracing Polynomial X-ray Optics:

In 1992, Burrows, Burg, and Giacconi showed how adding higher order polynomial terms to Wolter I prescriptions, and hence giving up some on-axis spatial resolution, can lead to mirror surface prescriptions with improved spatial resolution over a wide FOV. Optimizing a nested array of polynomial optics involves many Monte-Carlo ray traces to cover parameter space, and complex methods of finding the optimum design. We have devised a method, valid when the polynomial coefficients are sufficiently small, for ray tracing polynomial optics keeping the polynomial coefficients in symbolic form. The method treats a polynomial optic as a perturbation on an underlying Wolter I optic:

$$r_s^2(z) = r_{0,s}^2\left[1 + 2\,A_s\left(\frac{z}{r_{0,s}}\right) + B_s\left(\frac{z}{r_{0,s}}\right)^2 + u_{a,s}\left(\frac{z}{r_{0,s}}\right)^2 + u_{b,s}\left(\frac{z}{r_{0,s}}\right)^3\right]$$

where $u_{a,s}$ and $u_{b,s}$ are the polynomial coefficients and differ for the primary, *P*, and secondary, *S*, mirror elements (higher order polynomial terms may also be included), as do A_s and B_s.

For an underlying Wolter I optic we have:

$$P: \quad A_s = \tan\alpha_{0,s} \qquad B_s = 0$$

$$S: \quad A_s = \tan 3\alpha_0, s \qquad B_s = h(\alpha_{0,s})\tan^2 3\alpha_{0,s}$$

$$\alpha_{0,s} = \left(\frac{1}{4}\right)\tan^{-1}\left(\frac{r_{0,s}}{f}\right)$$

$$h(\alpha_{0,s}) = 1 - \frac{1}{[1+2\cos(2\alpha_{0,s})]^2}$$

The angle α0,s is the graze angle for an on-axis ray at the intersection plane.

Our method involves writing ray position components *(x,y,z)* and direction vectors *(k_x,k_y,k_z)* in the form:

$$\lambda = \lambda_{00} + \vec{u}\cdot\vec{\lambda}_0 + \vec{u}\cdot\overline{\overline{\lambda}}\cdot\vec{u} \qquad \text{where} \qquad \vec{u} = \left(u_{a,s,P},\ u_{b,s,P},\ u_{a,s,S},\ u_{b,s,S}\right)$$

In this notation, the mirror prescription becomes:

$$r_s^2(z) = r_{0,s}^2\left[1 + 2A_s\left(\frac{z}{r_{0,s}}\right) + B_s\left(\frac{z}{r_{0,s}}\right)^2 + \left(\frac{z}{r_{0,s}}\right)^2 \vec{u}\cdot\vec{\varsigma}_{0,s}\right]$$

$$\vec{\varsigma}_{0,s,P} = \left(1,\ \frac{z}{r_{0,s}},\ 0,\ 0\right) \qquad \vec{\varsigma}_{0,s,S} = \left(0,\ 0,\ 1,\ \frac{z}{r_{0,s}}\right)$$

As long as the polynomial coefficients are sufficiently small, various operations such as addition, subtraction, multiplication, division, taking of square roots, etc. are evaluated by expansions to second order in the polynomial coefficients. Then all the tasks required for a ray trace can be accomplished to the appropriate order. These tasks are (1) populating the entrance aperture with rays (in position, the incident direction is assumed), (2) finding intersections with mirror segment surfaces, (3) calculating unit normals to those surfaces, including deviations due to non-ideal surfaces, (4) determining the direction vector of the reflected ray, and (5) taking account of any obstruction by the next innermost shell.